



Europäisches
Patentamt

European
Patent Office

Rec. PCT/PTO
Office européen
des brevets

PCT/EP2004/541315
4/000845
30 JUN 2005

RECEIVED	
19 MAR 2004	
WIPO	PCT

Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-
gen stimmen mit der
ursprünglich eingereichten
Fassung der auf dem näch-
sten Blatt bezeichneten
europäischen Patentanmel-
dung überein.

The attached documents
are exact copies of the
European patent application
described on the following
page, as originally filed.

Les documents fixés à
cette attestation sont
conformes à la version
initialement déposée de
la demande de brevet
européen spécifiée à la
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03075207.5

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

PC/EP200 4 / 0 0 0 6 4 5

Anmeldung Nr.:
Application no.: 03075207.5
Demande no:

Anmeldetag:
Date of filing: 21.01.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

CropDesign N.V.
Technologiepark 3
9052 Zwijnaarde-Gent
BELGIQUE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se référer à la description.)

Regulatory sequence

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

C12N15/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT SE SI SK TR LI

072-GOS2-PROV

Regulatory sequence

The present invention relates to the field of plant molecular biology. In particular, the present invention describes the use of a regulatory promoter sequence of the rice GOS2 gene for the regulation of gene expression in plant cells of dicotyledonous plants.

Plant genetic engineering has the potential to enable the production of plants having all manner of desirable traits, such as increased pathogen resistance, herbicide resistance, increased yield, increased stress tolerance etc. This is typically achieved by inserting suitable pieces of DNA from a heterologous source, or from the same source, into a plant. The expression of heterologous DNA sequences in a plant host is dependent upon the presence of an operably linked promoter that is functional within the plant host.

Promoters are DNA elements that regulate transcription in prokaryotic and eukaryotic cells, and are located in the 5'-flanking or upstream sequence region of a transcribed gene. The sequence of a promoter varies in length and base pair composition from gene to gene. Within the sequence of a promoter, many regulatory sequence elements can be present. These regulatory sequence elements, which interact with DNA binding proteins that are part of the transcription initiation complex, are generally embedded in the promoter sequence. Specific elements, called DNA-boxes, contribute defined features or a defined activation pattern to a promoter. Many of these boxes are recognition sites to which regulatory transcription factors can bind and are part of a tight control mechanism of the promoter. Numerous plant promoter DNA boxes have been described, such as tissue-specific boxes (Chaubet *et al.*, Plant J. 10 (3) 425-435, 1996), pyrimidin boxes (Gubier & Jacobsen, Plant Cell 4 (11) 1435-1441, 1992) that influence the level of expression, or G-boxes (Dolferus *et al.*, Plant Physiol. 105 (4) 1075-1087, 1994), which reduce promoter activity upon exposure to cold or dehydration.

Recombinant DNA technology has now found a wide application in research and in product development. In molecular biology, specific promoters are frequently used for the expression of heterologous nucleic acid sequences or proteins in organisms, whereby the expression may be constitutive (i.e. continuous expression throughout the cells of an organism) or limited to defined parts of an organism (i.e. tissue or cell preferred expression) or to certain developmental stages or to certain environmental or physiological conditions (i.e. inducible expression). A need exists in various industrial applications for transcriptional control elements capable of driving gene-expression in plants. Therefore, in order to obtain desired gene expression, which can be fine tuned depending on specific needs, it is desirable to provide a wide array of promoters from which one can be chosen. Persons skilled in the art need to have

072-GOS2-PROV

at their disposal a variety of promoters with different expression patterns or with different time-control features from which to choose. For many industrial and agronomic applications, where the aim is to confer a given characteristic throughout the plant development and during any environmental or physical state of the plant, a constitutive promoter would be needed. A
5 constitutive promoter may be suitable for driving expression of selectable marker genes or may be used when the gene product of a particular gene is to be isolated and purified from the plant for commercial purposes.

Frequently it is also desirable to use promoters that are regulated in a similar way and have a
10 comparable expression pattern over a wide range of host species. The regulation of promoters, however, is often specific to the source organism. Promoters from one plant species may not work in another, even when they belong to the same family. It is thought that the more evolutionary distance there is between the source organism and the accepting organism, the less likely it is that the promoter from the source organism will be active and
15 properly regulated in the accepting organism.

Very few constitutive plant promoters that are active in a wide range of species have been isolated to date (Callis *et al.*, J. Biol. Chem. 265 12486-12493, 1990; Zhang *et al.*, Plant Cell 3 1155-1165, 1991; de Pater *et al.*, Plant J. 2 837-844, 1992; Mandel *et al.* Plant Mol. Biol. 29 995-1004, 1995; Baszczynski *et al.*, Maydica 42 189-201, 1997). As a result, the cauliflower
20 mosaic virus 35S (CaMV 35S; Odell *et al.*, Nature 313 653-660, 1985) is currently the most widely used constitutive promoter in plant science. Its regulation is well known and has been shown to work in many monocotyledonous and dicotyledonous plant species, and even in some fungal species. The main drawback of such a promoter however is its viral origin. Plants transformed with the CaMV35S promoter frequently show several undesirable characteristics,
25 such as efficient silencing of the transgenes or increased recombination, leading to instability over time and over subsequent generations (Kohli *et al.*, Plant J. 17 591-601, 1999; Al-Kaff *et al.*, Nature Biotechnology 18 995-999, 2000). In addition, viral elements are often not desirable in transgenic products, both for regulatory reasons and for reasons of general public acceptance.

30 The promoter of the rice GOS2 gene, which encodes a translation initiation factor, is a strong constitutive promoter and has been shown to be active in rice, barley and ryegrass and many monocotyledons, with expression levels similar to that of CaMV 35S.

35 It has now been found that the GOS2 promoter from rice is active in non-monocotyledonous plants, and has a similar expression pattern as the CaMV 35S promoter. Consequently, the present invention teaches the use of the GOS2 promoter for driving constitutive expression of

072-GOS2-PROV

nucleic acids in non-monocotyledonous plants. The present invention therefore provides a good alternative for the CaMV 35S promoter.

The promoter of the present invention, when linked to a useful gene and transferred in a plant or in a plant cell, enables an expression of that useful gene. In addition, since the promoter of the present invention is a constitutive promoter, it is possible to use it as a promoter to drive the expression of the transferred useful gene in all tissues at any developmental stage. Moreover, as the promoter of the present invention is a promoter originating from a monocotyledonous plant, it is less likely to be sensitive to silencing than other dicotyledonous promoters or the CaMV35S promoter, when used to express the transferred useful gene in dicotyledonous plants, such as *Arabidopsis thaliana*.

Those skilled in the art will be aware that the invention described herein is subject to variations and modifications other than those specifically described. It is to be understood that the invention described herein includes all such variations and modifications. The invention also includes all such steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any or more of these steps or features.

Throughout this specification, unless the context requires otherwise the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

As used herein, the term "derived from" shall be taken to indicate that a particular integer or group of integers has originated from the species specified, but has not necessarily been obtained directly from the specified source.

In accordance with the present invention, an isolated regulatory nucleic acid sequence is provided comprising a regulatory sequence as represented in SEQ ID NO 1 or a functional fragment or functional variant thereof, characterised in that said regulatory sequence is capable of driving expression of an associated nucleic acid sequence in a plant or plant cell, such plant or plant cell being of a systematic class other than Liliopsida. According to a preferred feature, the isolated regulatory nucleic acid sequence is a promoter sequence.

The terms "regulatory sequence" or "promoter" as used herein refer to DNA sequences which are necessary to effect expression of sequences to which they are ligated. The regulatory

072-GOS2-PROV

sequences may differ depending upon the intended host organism and upon the nature of the sequence to be expressed. In eukaryotes, regulatory sequences generally include promoters, terminators, and, in some instances, enhancers, and/or 5' and 3' untranslated sequences. The term "regulatory sequence" is intended to include, as a minimum, all components necessary for expression and may be also additional advantageous components. As used herein, a 'promoter' includes reference to a region of DNA upstream from the transcription start and involved in binding RNA polymerase and other proteins to start transcription. Reference herein to a "promoter" is to be taken in its broadest context and includes the transcriptional regulatory sequences derived from a classical eukaryotic genomic gene, including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence and additional regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or external stimuli, or in a tissue-specific manner. Consequently, a repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription is not specifically regulated, though it can vary under the influence of general metabolic conditions.

The term "promoter" also includes the transcriptional regulatory sequences of a classical prokaryotic gene, in which case it may include a -35 box sequence and/or a -10 box transcriptional regulatory sequences. The term "promoter" is also used to describe a synthetic or fusion molecule, or derivative which confers, activates or enhances expression of a nucleic acid molecule in a cell, tissue or organ. The present invention thus also encompasses fragments or variants of SEQ ID NO1, which may be joined by any means known in the art, e.g., by ligation *in vitro*, to create hybrid regulatory sequences comprising mixtures of parts of regulatory elements from different sources, either natural or synthetic. Therefore, the invention also provides a hybrid regulatory nucleic acid sequence comprising the regulatory sequence represented by SEQ ID NO 1 or a fragment or variant thereof, functionally combined with one or more other regulatory sequences or fragments thereof.

Promoters may contain additional copies of one or more specific regulatory elements, to further enhance expression and/or to alter the spatial expression and/or temporal expression of a nucleic acid molecule to which it is operably connected. Such regulatory elements may be placed adjacent to a heterologous promoter sequence to drive expression of a nucleic acid molecule in response to e.g. copper, glucocorticoids, dexamethasone, tetracycline, gibberellin, cAMP, abscisic acid, auxin, wounding, ethylene, jasmonate or salicylic acid, or to confer expression of a nucleic acid molecule to specific cells, tissues or organs such as meristems, leaves, roots, embryo, flowers, seeds or fruits. In the context of the present invention, the

072-GOS2-PROV

promoter preferably is a plant-expressible promoter sequence and more preferably the sequence represented in SEQ ID NO1, a functional fragment or a functional variant thereof. The term "functional", when used in respect of a nucleic acid or part thereof, is to be understood as having an essentially similar biological activity compared to the naturally occurring nucleic acid or part thereof.

The term "nucleic acid sequence(s)", "nucleic acid molecule(s)", "polynucleotide(s)", "nucleotide sequence(s)", "DNA sequence(s)" or "gene(s)", when used herein refers to nucleotides, either ribonucleotides (RNA) or deoxyribonucleotides (DNA) or a combination of both, in a polymeric form of any length. These terms furthermore include double-stranded and single-stranded DNA and RNA. These terms also include nucleotide modifications known in the art, such as methylation, cyclisation, 'caps', substitution of one or more of the naturally occurring nucleotides with an analogue such as inosine, and polynucleotide backbone modifications. The term "isolated" in the present application means removed from its original environment. For example, a nucleic acid present in the natural state in an organism is not isolated, whereas the same nucleic acid separated from the adjacent nucleic acids in which it is naturally present, is regarded as being "isolated".

A nucleic acid sequence may comprise coding sequences or it may contain non-coding sequences such as regulatory sequences. A "coding sequence" or "open reading frame" or "ORF" is defined as a nucleotide sequence that can be transcribed into mRNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences, i.e. when the coding sequence or ORF is present in an expressible format. The coding sequence or ORF is bound by a 5' translation start codon and a 3' translation stop codon. A coding sequence or ORF can include, but is not limited to RNA, mRNA, cDNA, recombinant nucleotide sequences, synthetically manufactured nucleotide sequences or genomic DNA. The coding sequence or ORF can be interrupted by intervening nucleic acid sequences.

Coding sequences, or genes essentially encoding the same protein but isolated from different sources, can consist of substantially divergent nucleic acid sequences. Reciprocally, substantially divergent nucleic acid sequences can be designed to effect expression of essentially the same protein. These nucleic acid sequences are the result of e.g. the existence of different alleles of a given gene, or of the degeneracy of the genetic code or of differences in codon usage. Thus, amino acids such as methionine and tryptophan are encoded by a single codon whereas other amino acids such as arginine, leucine and serine can each be translated from up to six different codons. Preferred codon usage of various organisms can be found in

072-GOS2-PROV

<http://www.kazusa.or.jp/codon>. Allelic variants are further defined as comprising single nucleotide polymorphisms (SNPs) as well as small insertion/deletion polymorphisms (INDELs; the size of INDELs is usually less than 100 bp). SNPs and INDELs form the largest set of sequence variants in naturally occurring polymorphic strains of most organisms.

5

The term "fragment of a sequence" or "part of a sequence" means a truncated sequence of the original sequence in question. The truncated sequence (nucleic acid or protein sequence) can vary widely in length; the minimum size being a sequence of sufficient size to provide a sequence with at least a comparable function and/or activity of the original sequence in question, while the maximum size is not critical. In some applications, the maximum size usually is not substantially greater than that required to provide the desired activity and/or function(s) of the original sequence. These terms are not restrictive to the content of the DNA fragment or segment, which can be any DNA, with any functionality. For example the DNA fragment or segments can comprise many genes, with or without additional control elements, it can contain just spacer sequences etc.

15

Furthermore, a variant of a DNA fragment encompasses a sequence showing homology to the sequence referred to is capable of promoting expression of an associated DNA sequence when reintroduced into a plant and is able to hybridise to the sequence referred to or to a portion thereof under stringent conditions. The person skilled in the art will appreciate that in this case, reactions under stringent conditions for hybridisation are typically carried out at a temperature of between 60°C and 65°C in 0.3 strength citrate buffer saline containing 0.1% SDS followed by rinsing at the same temperature with 0.3 strength citrate buffer saline containing 0.1% SDS.

20

"Expression" means the production of a protein or nucleotide sequence in a cell itself or in a cell-free system. It includes transcription into an RNA product, post-transcriptional modification and/or translation to a protein product or polypeptide from a DNA encoding that product, as well as possible post-translational modifications. A coding sequence can be transcribed in sense or antisense direction.

25

A "non-monocotyledonous plant" comprises all plant species which belong to the kingdom Plantae as defined by ITIS [the Integrated Taxonomic Information System (<http://www.itis.usda.gov/index.html>)], including the class of Magnoliopsida but without the class of the Liliopsida (Lillatae, Monocotyledonae), which are part of the division Magnoliophyta, which in its turn belongs to the subkingdom Tracheobionta.

30

072-GOS2-PROV

The isolated nucleic acid according to the present invention can be used to drive gene expression in any plant, more specifically a non-monocotyledonous plant, in particular it is applicable to a dicotyledonous plant including a fodder or forage legume, ornamental plant, food crop, tree, or shrub. The preferred host plant species is *Arabidopsis thaliana*.

5

Advantageously, the regulatory sequence represented by SEQ ID NO 1 or a functional fragment or functional variant thereof is capable of driving expression of an associated nucleic acid sequence in a non-monocotyledonous plant. Such an associated nucleic acid sequence may be any heterologous DNA sequence derived from any organism which sequence is to be expressed in a non-monocotyledonous plant. The heterologous DNA sequence may be obtained from any organism, including plants, which plant species may be the same plant species into which the sequence is to be introduced, or it may be different. Additionally or alternatively, the associated nucleic acid sequence may be endogenous nucleic acid sequence.

10

Therefore, the invention also provides an isolated regulatory nucleic acid sequence as defined above, wherein said associated nucleic acid sequence is an isolated nucleic acid sequence or a nucleic acid sequence endogenous to the host cell in which the promoter is introduced.

15

The term "associated" as used herein, or "operably linked" refers to a juxtaposition between a regulatory and a coding sequence, wherein the components so described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences. In case the regulatory sequence is a promoter, it would be known to a skilled person that double-stranded nucleic acid is preferable. The associated nucleic acid sequence encompasses heterologous nucleic acids. Heterologous nucleic acids refer to endogenous nucleic acids, i.e. nucleic acids that originate from within the cell but that are not naturally located in the cell, or are located in a different chromosomal site of the cell. Heterologous nucleic acids can also be derived from another species and may be introduced as a transgene for example by transformation. This transgene may be substantially modified from its native form in composition and/or genomic environment through deliberate human manipulation. Also expression of the native nucleic acid sequences can be modified by introduction in the plant of regulatory sequences according to the invention capable of altering expression of the native nucleic acid sequences. Native nucleic acid sequences refer to nucleic acid sequences in their natural location in the genome of an organism.

25

30

35

In a second embodiment, there is provided a recombinant nucleic acid, comprising at least:

072-GOS2-PROV

- (i) a first nucleic acid sequence comprising a regulatory sequence as represented in SEQ ID NO 1 or a functional fragment or functional variant thereof, characterised in that said regulatory sequence is capable of driving expression of an associated nucleic acid sequence in a plant or plant cell, such plant or plant cell being of a systematic class other than Liliopsida; and optionally,
- (ii) an associated nucleic acid sequence, wherein the associated nucleic acid sequence is heterologous to the first nucleic acid.

The recombinant nucleic acid encompasses any means for the cloning of and/or transfer of a nucleic acid into a host cell, including "vectors" or "vector sequences", which can be introduced in an organism by transformation or transfection, and which may either be integrated into the genome of the host cell or maintained in some form extrachromosomally. Vector sequences generally comprise a set of unique sites recognised by restriction enzymes, the multiple cloning site (MCS), wherein one or more non-vector sequence(s) can be inserted.

With "non-vector sequence" or "insert" is accordingly meant the desired DNA segment which one wishes to clone in one or more of the sites of the MCS, comprised within a vector. The insert can have one or more genes.

"Expression vectors" form a subset of vectors, which, by virtue of comprising the appropriate regulatory sequences enabling the creation of an expressible format for the inserted non-vector sequence(s), thus allowing transcription and/or translation of a nucleic acid inserted therein. Expression vectors are known in the art enabling protein expression in organisms including bacteria (e.g. *E. coli*), fungi (e.g. *S. cerevisiae*, *S. pombe*, *Pichia pastoris*), insect cells (e.g. baculoviral expression vectors), animal cells (e.g. COS or CHO cells) and plant cells (e.g. potato virus X-based expression vectors, see e.g. Vance *et al.* 1998 - WO9844097). Expression vectors can for instance be cloning vectors, binary vectors or integrating vectors.

Regulatory elements ensuring expression in prokaryotic and/or eukaryotic cells are well known to those skilled in the art. A preferred regulatory sequence is a plant regulatory sequence, more preferably the regulatory sequence listed in SEQ ID NO1, a functional fragment or a functional variant thereof. When used in eukaryotic cells vectors may also comprise terminators which contain polyadenylation signals ensuring 3' processing and polyadenylation of a primary transcript and termination of transcription. In plant cells the termination signals usually employed are from the *nopaline synthase* gene or the CaMV 35S terminator. Additional regulatory elements may include transcriptional as well as translational enhancers. A plant translational enhancer often used is the CaMV omega sequences. The inclusion of an intron

072-GOS2-PROV

has been shown to increase expression levels by up to 100-fold in certain plants (Maiti *et al.*, Transgenic Research 6 143-156, 1997; Ni, Plant Journal 7 (1995), 661-676).

- Advantageously, vectors of the invention comprise a selectable and/or scorable marker.
- 5 Selectable marker genes useful for the selection of transformed plant cells, callus, plant tissue and plants are well known to those skilled in the art. Antimetabolite resistance is very useful as a basis for selection: for example, the *dhfr* gene confers resistance to methotrexate (Reiss *et al.*, Plant Physiol. (Life Sci. Adv.) 13 143-149, 1994); the *npt* gene confers resistance to the aminoglycosides neomycin, kanamycin and paromomycin (Herrera-Estrella *et al.*, EMBO J. 2 987-995, 1983); or the *hpt* gene, which confers resistance to hygromycin (Marsh, Gene 32 481-485, 1984). Additional selectable markers genes have been described, such as *trpB*, which allows cells to utilize indole in place of tryptophan; *hisD*, allowing cells to utilize histinol in place of histidine (Hartman, Proc. Natl. Acad. Sci. USA 85 8047, 1988); *mannose-6-phosphate isomerase* enabling cells to utilize mannose (WO 94/20627) and *ornithine decarboxylase* which confers resistance to the ornithine decarboxylase inhibitor, 2-
- 10 (difluoromethyl)-DL-ornithine or DFMO (McConlogue, 1987, in: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed.) or the *deaminase* gene from *Aspergillus terreus* which confers resistance to Blasticidin S (Tamura, Biosci. Biotechnol. Biochem. 59 2336-2338, 1995). Useful scorable markers are also known to those skilled in the art and are commercially available. Advantageously, the marker is a gene encoding luciferase
- 20 (Giacomini, P.I. Sci. 116 59-72, 1996; Srikantha, J. Bact. 178 121, 1996), green fluorescent protein (Gerdes, FEBS Lett. 389 44-47, 1996; Haseloff *et al.*, Proc. Natl. Acad. Sci. U.S.A. 94 2122-2127, 1997) or β -glucuronidase (Jefferson, EMBO J. 6 3901-3907, 1987). Vectors can further provide primer sites, e.g., for PCR, transcriptional and/or translational initiation and/or
- 25 regulation sites, recombinational sites, replicons, etc.
- The current invention clearly includes any recombinant nucleic acid, vector or expression vector comprising the regulatory sequence according to the present invention and/or a non-vector sequence as defined supra.

- 30 According to another embodiment, the invention relates to a non-monocotyledonous plant cell comprising a recombinant nucleic acid as described above. The recombinant nucleic acid may be stably integrated into the genome of the non-monocotyledonous plant cell. The invention also provides a callus or a plant derived from these plant cells. Also within the scope of the present invention is a harvestable part, organ, tissue or propagation material of a plant
- 35 according to the invention, comprising a nucleic acid sequence comprising a regulatory sequence represented by SEQ ID NO: 1 or a fragment or variant thereof.

072-GOS2-PROV

Furthermore, there is provided a method for expression of a nucleic acid sequence in a plant cell of a systematic class other than Liliopsida, which method comprises introducing into a plant cell a regulatory sequence represented by SEQ ID NO 1 or a functional fragment or functional variant thereof, wherein said regulatory sequence is capable of driving expression of said nucleic acid sequence which is either an isolated or endogenous nucleic acid sequence.

Means for introducing recombinant DNA into plant tissue or cells include, but are not limited to, transformation using CaCl_2 and variations thereof, in particular the method described previously (Hanahan, J. Mol. Biol. 166 (4) 557-580, 1983)), direct DNA uptake into protoplasts (Krens *et al.*, Nature 296 72-74, 1982; Paszkowski *et al.*, EMBO J. 3 2717-2722, 1984), PEG-mediated uptake to protoplasts (Armstrong *et al.*, Plant Cell Reports 9 335-339, 1990) microparticle bombardment, electroporation (Fromm *et al.*, Proc. Natl. Acad. Sci. U.S.A. 82 5824-5826, 1985) microinjection of DNA (Crossway *et al.*, Mol. Gen. Gen. 202 179-185, 1986; Fromm *et al.* Proc. Natl. Acad. Sci. U.S.A. 82 5824-5826, 1985), microparticle bombardment of tissue explants or cells (Christou *et al.*, Plant Physiol. 87 671-674, 1988), vacuum-infiltration of tissue with nucleic acid, or in the case of plants, T-DNA-mediated transfer from *Agrobacterium* to the plant tissue as described essentially (An *et al.*, EMBO J. 4 277-284, 1985; Dodds, Plant genetic engineering, 1985; Herrera-Estrella *et al.*, EMBO J. 2 987-995, 1983; Herrera-Estrella *et al.*, Nature 303 209-213, 1983). Methods for transformation of monocotyledonous plants are well known in the art and include *Agrobacterium*-mediated transformation (Cheng *et al.* 1997 - WO9748814; Hansen 1998 - WO9854961, Hiei *et al.* 1994 - WO9400977; Hiei *et al.* 1998 - WO9817813; Rikilshi *et al.* 1999 - WO9904618; Saito *et al.* 1995 - WO9506722), microprojectile bombardment (Adams *et al.* 1999 - US5969213; Bowen *et al.* 1998 - US5736369; Chang *et al.* 1994 - WO9413822; Lundquist *et al.* 1999 - US5874265/US5990390; Vasil and Vasil 1995 - US5405765; Walker *et al.* 1999 - US5955362), DNA uptake (Eyal *et al.* 1993 - WO9318168), microinjection of *Agrobacterium* cells (von Holt 1994 - DE4309203) and sonication (Flner *et al.* 1997 - US5693512).

For microparticle bombardment of cells, a microparticle is propelled into a cell to produce a transformed cell. Any suitable ballistic cell transformation methodology and apparatus can be used in performing the present invention. Exemplary apparatus and procedures are disclosed by Stomp *et al.* (U.S. Patent No. 5122466) and Sanford and Wolf (U.S. Patent No. 4945050). When using ballistic transformation procedures, the gene construct may incorporate a plasmid capable of replicating in the cell to be transformed.

Examples of microparticles suitable for use in such systems include 1 to 5 μm gold spheres. The DNA construct may be deposited on the microparticle by any suitable technique, such as by precipitation.

072-GOS2-PROV

The nucleic acid molecule of the invention, or a genetic construct comprising it may be introduced into a cell using any known method for the transfection or transformation.

- 5 The invention further relates to a method for the production of transgenic plants, plant cells or plant tissues comprising the introduction in a plant, plant cell or plant tissue of a first nucleic acid molecule according to the invention operably linked to an associated nucleic acid sequence, wherein the associated nucleic acid sequence is heterologous to the first nucleic acid. According to another preferred embodiment, the invention relates to a transgenic plant
- 10 cell or plant as described here above wherein the nucleic acid of the invention is stably integrated into the genome of the plant cell or plant.

- Wherein a cell is transformed by the genetic construct of the invention, a whole plant may be regenerated from the transformed or transfected cell, in accordance with procedures well
- 15 known in the art. Plant tissue capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with a gene construct of the present invention and a whole plant regenerated therefrom. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular species being transformed. Exemplary tissue targets include leaf disks, pollen, embryos,
- 20 cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g., apical meristem, axillary buds, and root meristems), and induced meristem tissue (e.g., cotyledon meristem and hypocotyl meristem). The term "organogenesis", as used herein, means a process by which shoots and roots are developed sequentially from meristematic centres. The term "embryogenesis", as used herein, means a process by which shoots and
- 25 roots develop together in a concerted fashion (not sequentially), whether from somatic cells or gametes.

- The generated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1)
- 30 transformed plant may be selfed to give homozygous second generation (or T2) transformants, and the T2 plants further propagated through classical breeding techniques.
- The generated transformed plants contemplated herein may take a variety of forms. For example, they may be chimeras of transformed cells and non-transformed cells; clonal transformants (e.g., all cells transformed to contain the expression cassette); grafts of
- 35 transformed and untransformed tissues (e.g., in plants, a transformed rootstock grafted to an untransformed scion).

072-GOS2-PROV

The invention not only relates to a transgenic plant or plant tissue comprising plant cells as herein described but extends also to the harvestable parts of the transgenic plant, preferably selected from the group consisting of seeds, leaves, flowers, fruits, stem cultures, roots, tubers, rhizomes and bulbs. The invention also relates to the progeny derived from any of these transgenic plants or plant parts.

The present invention also encompasses any use of an isolated nucleic acid sequence comprising a regulatory sequence as represented by SEQ ID NO 1 or a functional fragment or functional variant thereof, characterised as above; including the use for driving constitutive expression of an associated sequence in a plant cell or plant of a systematic class other than Liliopsida. Also included is the use of a hybrid regulatory nucleic acid as defined above for driving expression of an associated nucleic acid sequence.

Description of figures

- Fig 1:** Typical GUS expression pattern in a plant transformed with a p2203 construct and a control plant. The GOS2 promoter shows a strong constitutive expression in shoot tissues, including the petioles, the vascular system, trichomes and in root tips. Expression in other root tissues is lower, but still easily noticeable.
- Fig 2:** GUS expression under control of the GOS2 promoter (lines AT0476-09 and AT0476-11), compared to GUS expression under control of the 35S promoter (line WS35S:GUS). For each line, 5 replicas were tested. The plants have 2 cotyledons and 2 true leaves and were stained for 1 hour (panel a) or were stained overnight (panel b).
- Fig 3:** GUS expression under control of the GOS2 promoter (lines AT0476-09 and AT0476-11), compared to GUS expression under control of the 35S promoter (line WS35S:GUS). For each line, 5 replicas were tested. The plants are in the 10 leaves stage and were stained for 1 hour (panel a) or were stained overnight (panel b).
- Fig 4:** GUS expression under control of the GOS2 promoter (lines AT0476-09 and AT0476-11), compared to GUS expression under control of the 35S promoter in line WS35S:GUS. For each line, 5 replicas were tested. The plants are in the adult stage with an inflorescence and were stained for 1 hour (panel a) or were stained overnight (panel b).

072-GOS2-PROV

Examples

The present invention will now be illustrated with the following examples.

DNA manipulation

- 5 All DNA procedures were performed according to standard protocols (Maniatis T *et al.* (2001). Molecular Cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, CSH, New York).

Example 1: Cloning of GUS with intron coding sequence

- 10 The *E. coli* β -glucuronidase gene, interrupted by the second Intron of the Potato light-inducible tissue-specific ST-LS1 gene, was amplified from vector pTHW136 (with a cassette identical to the one of pMOG563, MOGEN international n.v, LEIDEN, NETHERLANDS). The cDNA was amplified by PCR with Platinum Pfx DNA polymerase (Invitrogen) and using a sense primer including attB1 (SEQ ID NO2) and an antisense primer including attB2 (SEQ ID NO3).

- 15 Conditions for PCR were: 2 minutes of denaturation at 94°C (1 cycle); 35 cycles of 1 minute denaturation at 94°C, 1 minute annealing at 58°C and 2 minutes amplification at 68°C, and finally 1 cycle of 5 minutes elongation at 68°C. A prominent fragment with the expected size of 2 kb was isolated from gel and purified using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Orange, California).

- 20 The purified PCR fragment was used in a standard Gateway™ BP reaction (Invitrogen) with pDONR201 as a recipient vector. The identity and base pair composition of the insert was confirmed by sequencing. The integrity of the resulting plasmid was verified using restriction digests and was given the designation p0604 (entry clone). pDONR201 was obtained from

- 25 The purified PCR fragment was used in a standard Gateway™ BP reaction (Invitrogen) with pDONR201 as a recipient vector. The identity and base pair composition of the insert was confirmed by sequencing. The integrity of the resulting plasmid was verified using restriction digests and was given the designation p0604 (entry clone). pDONR201 was obtained from
- 30 p0604 is a Gateway™ "entry clone", and was used as such in a standard Gateway™ LR reaction, with p0640 as "destination vector". The resulting vector was p02203, its integrity was also checked by restriction digest analysis. The vector, containing as functional elements within the T-DNA region a selectable marker gene and a "Gateway cassette" intended for LR cloning of sequences of interest, was used for the transformation of *Arabidopsis thaliana*. Expression of these sequences of interest, upon being recombined into p0640, was driven by the regulatory sequence as listed in SEQ ID NO1.

072-GOS2-PROV

Example 2: Transformation of plants**5 Cultivation of the parental plants**

For the parental plants, approximately 12mg of wild-type *Arabidopsis thaliana* (ecotype Columbia) seeds were suspended in 27.5 ml of 0.2 % agar solution. The seeds were incubated for 2 to 3 days at a temperature of 4°C and then sown. The plants were germinated under standardised conditions: 22°C during the day, 18°C at night, with a relative humidity of 65-70%, 12 hours of photoperiod and sub-irrigation with water for 15 min every 2 or 3 days. The seedlings were then transplanted into pots with a diameter of 5.5cm, containing a mixture of sand and peat in a ratio of 1 to 3. The plants were then further grown under the same standard conditions as mentioned above.

15 Agrobacterium growth conditions and preparation

Agrobacterium strain C58C1RIF with helper plasmid pMP90 and vector p2203 was inoculated in a 50 ml plastic tube containing 1 ml LB (Luria-Bertani Broth) without antibiotic. The culture was shaken for 8-9h at 28°C. Next, 10 ml of LB without antibiotic was added to the plastic tube and shaken overnight at 28°C. Following this, the OD at 600 nm was measured. At an optical density of approximately 2.0, 40 ml of 10% sucrose and 0.05% Silwet L-77 (a mixture of 84 % polyalkyleneoxide modified heptamethyltrisiloxane and 16% allyloxypolyethyleneglycol methyl ether, OSI Specialties Inc) was added to the culture. The *Agrobacterium* culture thus obtained was used to transform the grown plants.

25 Flower dip

When each parental flower had one inflorescence of 7-10 cm in height, the inflorescences were submerged in the *Agrobacterium* culture and swirled gently for 2-3 seconds. For each transformation 2 plants were used. Hereafter, the plants were returned to the growing conditions as described above.

30

Seed collection

5 weeks after the flowers were dipped in the *Agrobacterium* culture, watering of the plants was stopped. The plants were further incubated at 25°C with a photoperiod of 20 hours. One week later, the seeds were harvested and placed in a seed drier during one week. The seeds were then cleaned, collected in 15 ml plastic tubes and stored at 4°C until further processing.

072-GOS2-PROV**T1 seeds generation**

Transgenic T0 seeds were selected under UV light for their GFP expression, and germinated. Plants were grown as described above, until T1 seeds were harvested. The different lines were given the name AT0476-xx, where xx is a number.

5

Example 3: Comparison of the GOS2 and the 35S promoter in Arabidopsis:

The expression pattern of the β -glucuronidase gene and strength of the GOS2 promoter were compared with those of the 35S promoter in Arabidopsis.

10 **Plant material used:**

AT0476-11: cassette GOS2-GUS and cassette 35S GFP (vector p2203) (T1 seeds)

AT0476-09: cassette GOS2-GUS and cassette 35S GFP (vector p2203) (T1 seeds)

WS35S:GUS: cassette 35S-GUS, no GFP selection, homozygous line from Versailles (T2 homozygous seeds), Wassilewskija (WS) background.

15

Procedure:

Around 100 seeds, after being kept at a temperature of 4°C for at least 2 days, were sown in soil and the plants were cultivated under standard conditions: 22°C during the day, 18°C at night, 65-70% relative humidity, 12 hours of photoperiod, sub-irrigation with water for 15 min every 2 or 3 days.

20

When the plants had 2 cotyledons and 2 true leaves, GFP expression in AT0476 lines was monitored under UV light and plants with a uniform GFP expression pattern were retained.

Next, 10 plants of each GOS2 or 35S lines were carefully isolated from the soil and labelled for identification. 5 of the plants were stained for GUS during 1 hour, the other 5 plants were

25

stained overnight. The short staining time is needed to assess qualitatively the relative strength of the promoters. If the staining reaction is saturated in less than 1 hour, the incubation time must be reduced. If the staining is very faint blue after one hour, the incubation time may be

increased to 2-3 hours. Finally, the stained plants were photographed. For each line, 20 other plants were carefully isolated from the soil and transferred in individual small pots for further

30

cultivation. When the plants had reached the 10-leave stage, again 10 plants of each line were isolated, labelled and stained as described above. The stained plants were photographed. At flowering time, 10 plants per line with siliques in various maturation stages were isolated and processed as above.

35 **Plant GUS staining procedure**

The material was covered with 90 % ice-cold acetone and incubated for 30 min at 4°C. After 3 washes of 5 minutes with Tris buffer [15,76 g Trizma HCl (Sigma T3253) + 2,922 g NaCl in 1 l

072-GOS2-PROV

double distilled water, adjusted pH to 7,0 with NaOH], the material was immersed in a Tris/ferricyanate/X-Gluc solution [9,8 ml Tris buffer + 0,2 ml ferricyanate stock (0,33 g Potassium ferricyanate (Sigma P3667) in 10 ml Tris buffer)+ 0,2 ml X-Gluc stock (26,1 mg X-Gluc (Europa Bioproducts ML 113A) in 500 µl DMSO)]. Vacuum infiltration was applied for 15 to 30 minutes. Next, the samples were incubated for an appropriate time (up to 16 hours) at 37 °C for development of the blue colour. The samples were washed 3 times during 5 minutes with Tris buffer. Chlorophyll was extracted in a series of 50 %, 70 % and 90 % ethanol (each for 30 minutes) with refreshments of the 90 % solution if necessary.

10 Results

The GOS2 promoter shows a strong constitutive expression in shoot tissues, including the petioles, the vascular system and trichomes. Except for the root tips, the expression in root tissues is lower, but is still easily detectable (Fig 1).

For all the plants with the same incubation time during staining (1hr or overnight), the pattern of GUS expression was similar. The GOS2 staining was as strong as 35S, at all stages that were sampled (2 leaves+ 2 cotyledons, 10 leaves stage and flowering stage). GOS2 appears equally strong and equally constitutive as 35S in *Arabidopsis* (Figs 2, 3 and 4).

072-GOS2-PROV

Claims

- 1) An isolated regulatory nucleic acid sequence comprising a regulatory sequence as represented in SEQ ID NO 1 or a functional fragment or functional variant thereof, characterised in that said regulatory sequence is capable of driving expression of an associated nucleic acid sequence in a plant or plant cell, said plant or plant cell being of a systematic class other than Liliopsida.
- 2) An isolated regulatory nucleic acid sequence according to claim 1, wherein said associated nucleic acid sequence is an isolated nucleic acid sequence or a nucleic acid sequence endogenous to the host cell in which the promoter is introduced.
- 3) An isolated regulatory nucleic acid sequence according to claim 1, which sequence is a promoter sequence.
- 4) A hybrid regulatory nucleic acid sequence comprising the regulatory sequence represented by SEQ ID NO 1 or a fragment or variant thereof, functionally combined with one or more other regulatory sequences or fragments thereof.
- 5) A recombinant nucleic acid, comprising at least:
 - (i) an isolated regulatory nucleic acid sequence according to claim 1; and optionally,
 - (ii) an associated nucleic acid sequence, wherein said associated nucleic acid sequence is heterologous to the sequence of (i).
- 6) A plant cell being of a systematic class other than Liliopsida comprising or having stably integrated into its genome a recombinant nucleic acid according to claim 5.
- 7) A callus or a plant consisting essentially of plant cells according to claim 6.
- 8) A harvestable part, organ, tissue or propagation material of a plant according to claim 7.
- 9) Method for expression of a nucleic acid sequence in a plant or plant cell of a systematic class other than Liliopsida, said method comprising introducing into said plant or plant cell a regulatory sequence represented by SEQ ID NO 1 or a functional fragment or functional

072-GOS2-PROV

variant thereof, wherein said regulatory sequence is capable of driving expression of said nucleic acid sequence which is either an isolated or endogenous nucleic acid sequence.

10) Use of an isolated nucleic acid sequence according to claim 1 for driving expression of
5 an associated nucleic acid sequence in a plant or plant cell, said plant or plant cell being of a systematic class other than Liliopsida.

11) Use of the hybrid regulatory nucleic acid according to claim 4 for driving expression of
an associated nucleic acid sequence.

10

072-GOS2-PROV

Abstract

Regulatory sequence

5

The present invention relates to the field of plant molecular biology. In particular, it describes the use of a regulatory nucleic acid sequence of the rice gene GOS2 for the regulation of gene expression in plant cells derived from plants other than monocotyledonous plants. The use of the regulatory sequence of the present invention results in constitutive expression with expression levels similar to that of CaMV 35S.

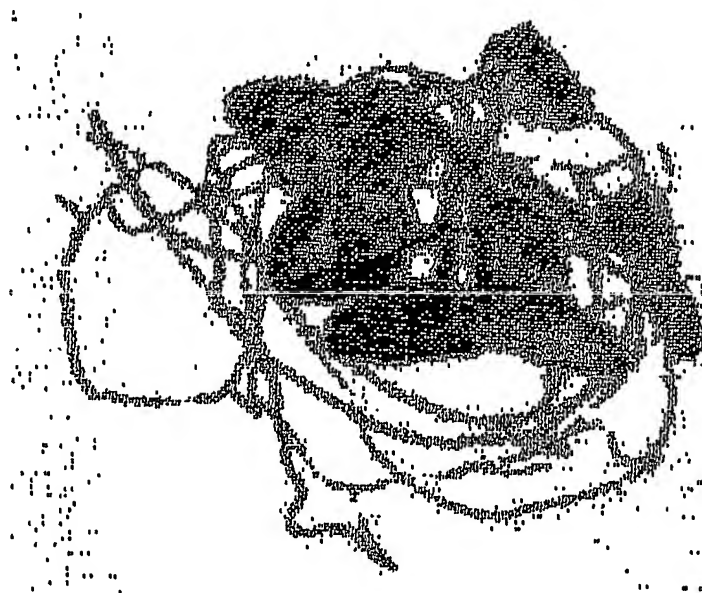
10

The present invention also relates to vectors and host cells comprising these nucleic acid sequences. The invention further relates to transgenic cells and plants comprising these sequences and to methods for obtaining such cells and plants.

072-GOS2-PROV

1/5

A



B



FIGURE 1

072-GOS2-PROV

2/5

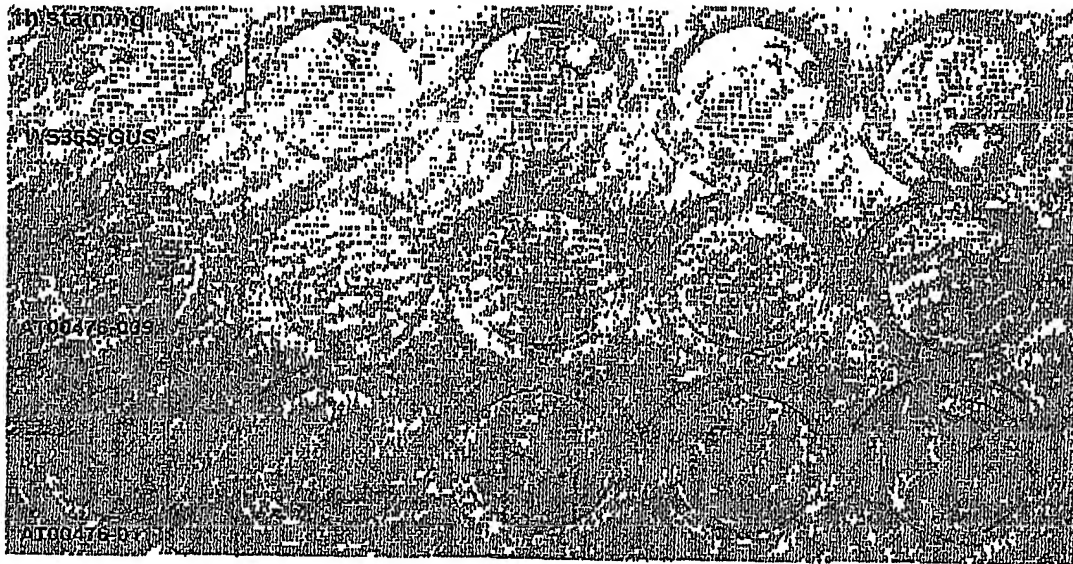


FIGURE 2

072-GOS2-PROV

3/5

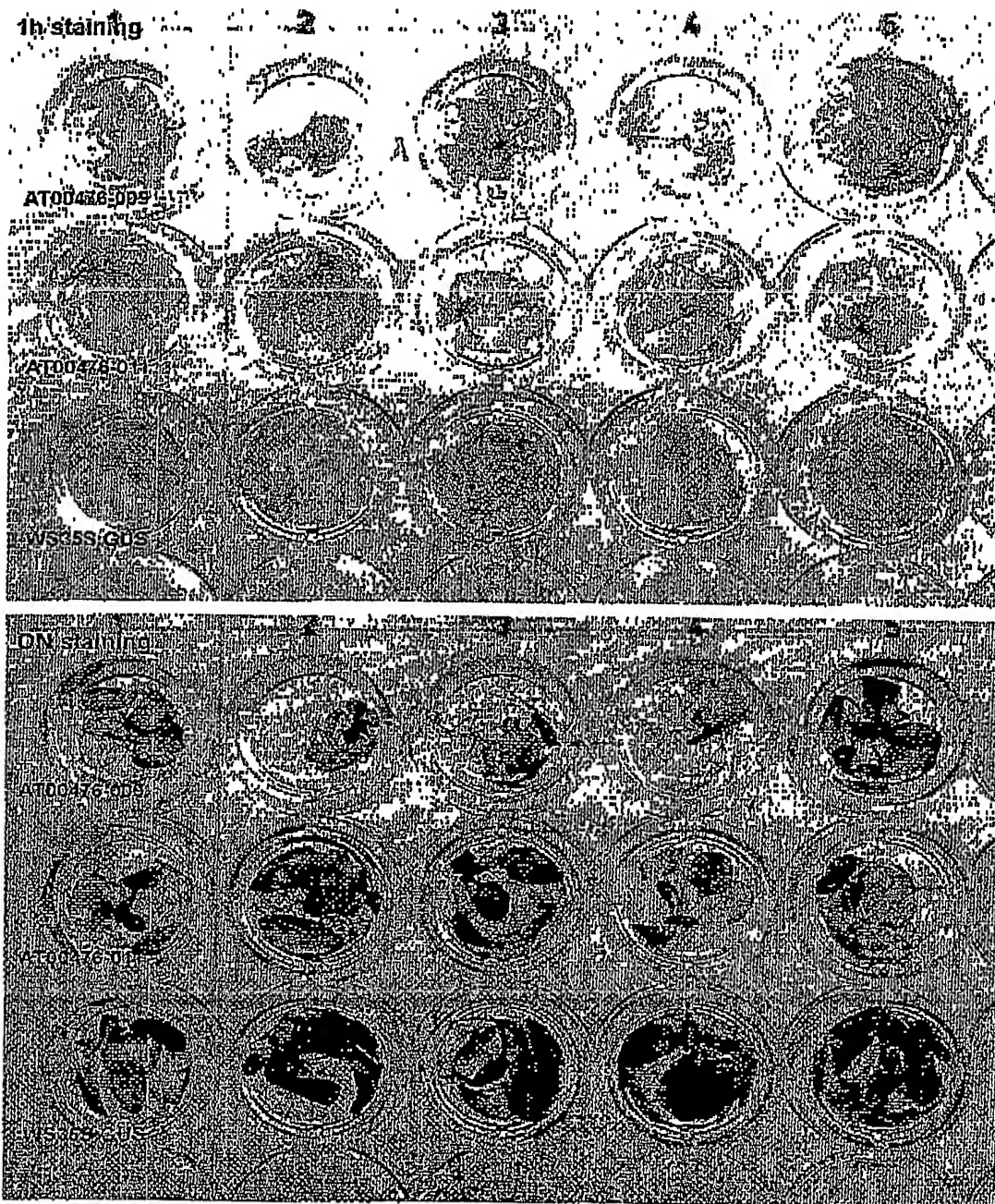


FIGURE 3

072-GOS2-PROV

4/5

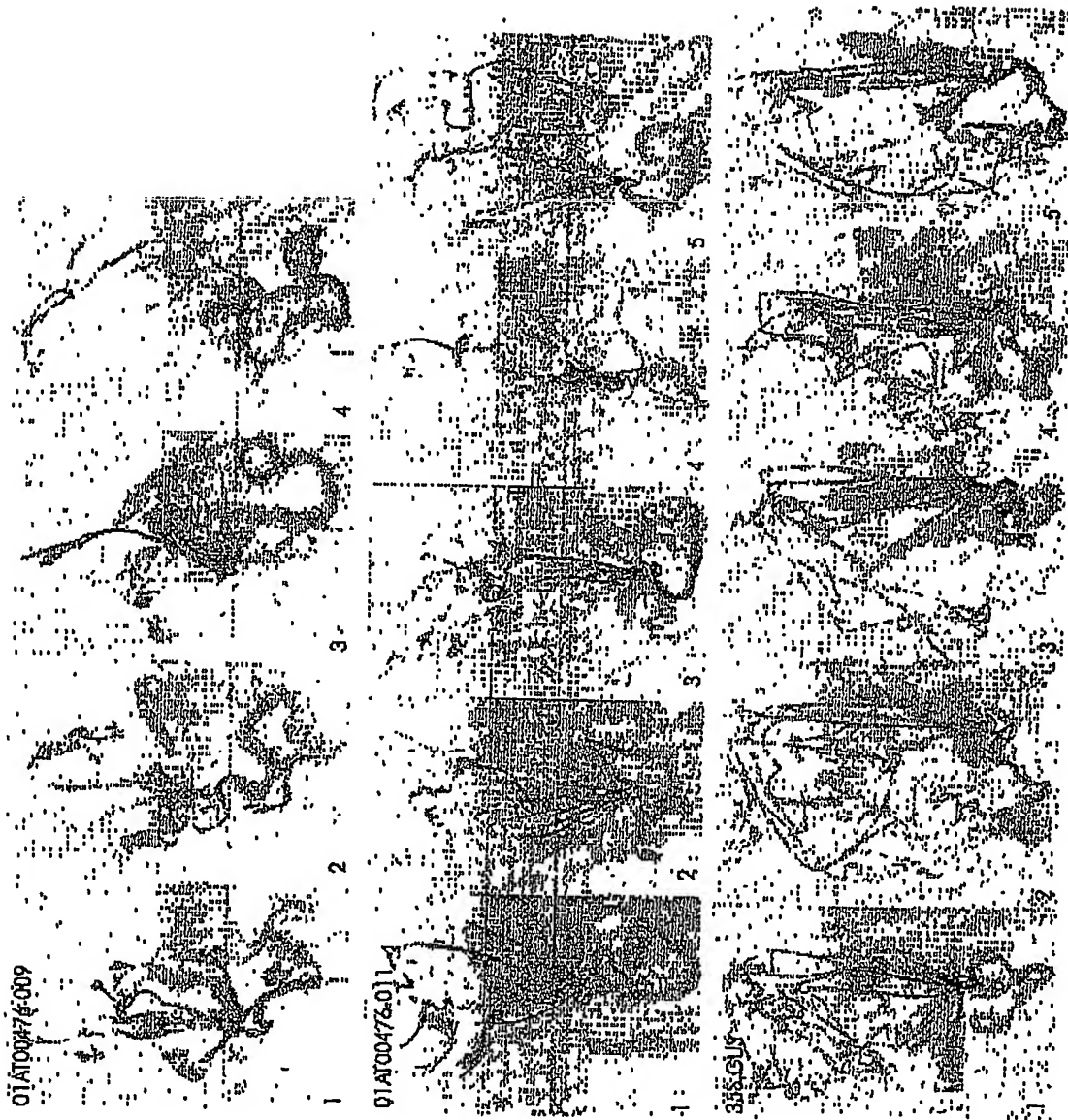


FIGURE 4

072-GOS2-PROV

5/5

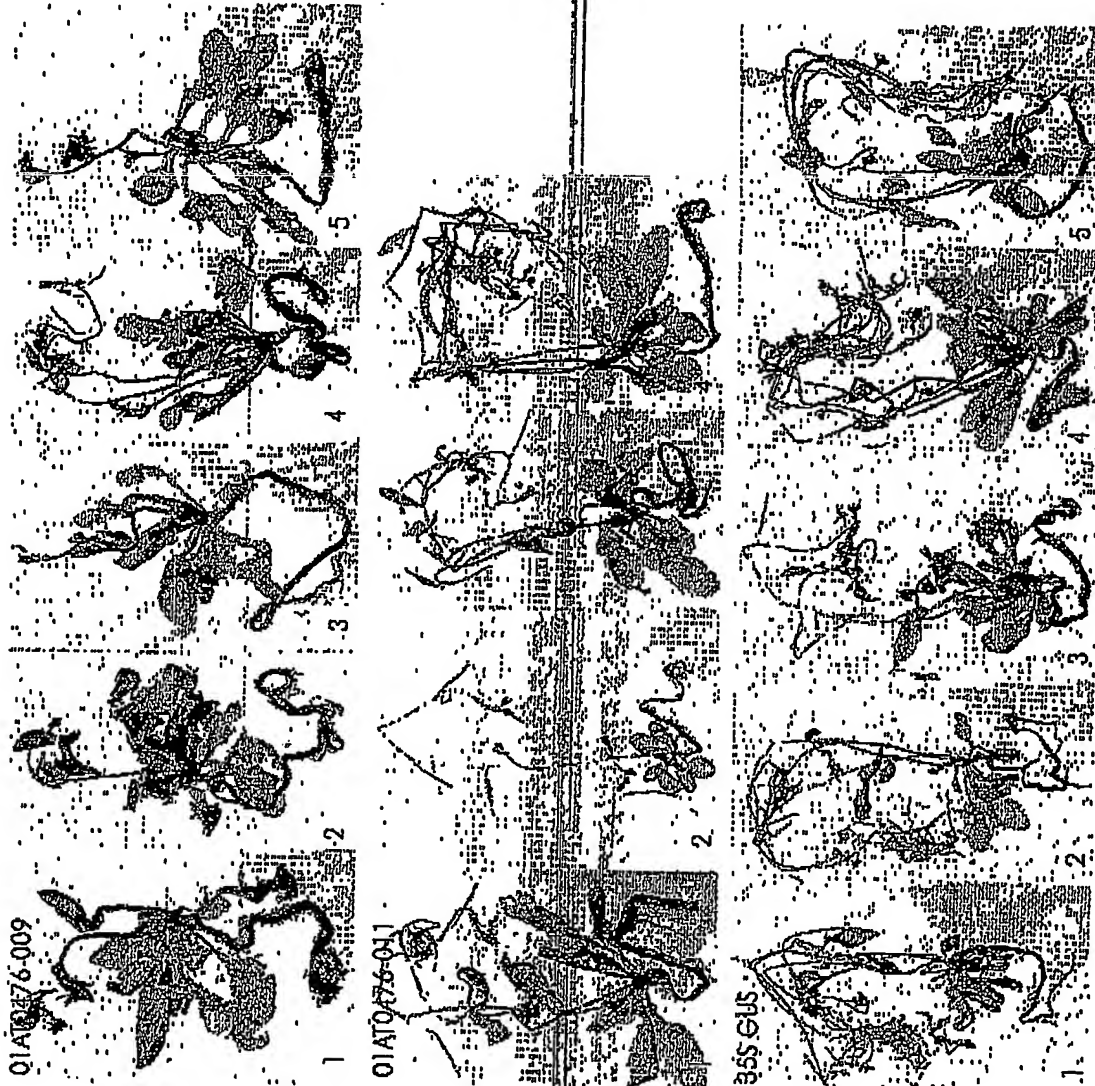


FIGURE 4 (continued)

072-GOS2-PROV.ST25.txt
SEQUENCE LISTING

<110> CropDesign N.V.

<120> Regulatory sequence

<130> 072-GOS2-PROV

<160> 3

<170> PatentIn version 3.1

<210> 1

<211> 2195

<212> DNA

<213> Oryza sativa

<400> 1

aatccgaaaa gtttctgcac cgttttcacg tcctaactaa caatataggg aacgtgtgct	60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgt aagaaaaact	120
catccacctt ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtcttgt	180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc	240
tctgtcatga agttaaatga ttccaggttag ccataattgt catcaaaactc ttcttgaata	300
aaaaaatott tctagctgaa ctcaatgggt aaagagagat attttttttt aaaaaaaat	360
agaatgaaga tattctgaac gtatcggcaa agatttaaac atataattat ataattttat	420
agtttgtgca ttcgttatat cgcacgtcat taaggacatg tcttactcca tctcaatttt	480
tatttagtaa ttaaagacaa ttgaattatt ttattatttt atcttttttc gattagatgc	540
aaggtaacta cgcacacact ttgtgctcat gtgcatgtgt gagtgcacct cctcatacac	600
gttcaactag cgacacatct ccaatatcac tcgctatttt aatacattta ggtagcaata	660
tctgaattca agcacttcac catcaccaga ccacttttaa taatatctaa aatacaaaaa	720
ataattttac agaatagcat gaaaagtatg aaacgaacaa tttagggtttt tcacatacaa	780
aaaaaaaaag aatttttgctc gtgcgcgagc gccaatcttc catattgggc acacaggcaa	840
caacagagtg gctgccaca gaacaaccca caaaaaacga tgatctaacy gaggaacagca	900

072-GOS2-PROV ST25.txt

```

agtcgcgaac aaccttttaa cagcaggctt tgcggccagg agagaggagg agaggcaaag 960
aaaaccaagc atcctcctcc tcccatctat aaattcctc ccccttttcc cctctctata 1020
taggaggcat ccaagccaag aagagggaga gcaccaagga caccgcacta gcagaagcog 1080
agcgacgcgc ttcttcgata catatcttcc ggtcgagttc ttggtcgata tcttccctcc 1140
tccacctcct cctcacaggg tatgtgccct tgggttctc ttggatttat tgttctaggt 1200
tgtgtagtag gggcggtgat gttaggaaag gggatctga totgtgatga ttcctgttct 1260
tggatttggg atagaggggt tcttgatgtt gcatgtctc ggttcgggtt gatttagtag 1320
atgggtttca atcgtctgga gagctctatg gaaatgaaat ggtttagggt acggaatctt 1380
gcgattttgt gagtacottt tgtttgaggt aaaatcagag caccgggtgat tttgcttggg 1440
gtaataaaaag tacatttgtt tggtcctcga ttctggtagt gatgottctc gatttgacga 1500
agctatcctt tgtttattcc ctattgaaca aaaataatcc aactttgaag acggtcccggt 1560
tgatgagatt gaatgattga ttcttaagcc tgcctaaat ttogoagctg gcttggttag 1620
atacagtagt ccccatcacg aaattcatga aaacagttat aatcctcagg aacaggggat 1680
tccctgttct tccgatttgc tttagtccca gaatttttt tcccaaatat cttaaaaagt 1740
cactttctgg ttcagttcaa tgaattgatt gctacaaat atgcttttat agcgttatcc 1800
tagctgtagt tcagtttata ggtaataccc ctatagttca gtcaggagaa gaacttatcc 1860
gattttctgat ctccattttt aatttatga aatgaactt agcataagca gtattcattt 1920
ggattatttt ttttatttagc tttcaccctt tcattattt gagctgaaag tctggcatga 1980
actgtcctca attttgtttt caaattcaca tgcattatt atcgattatc ctcttgtagt 2040
tacctgtaga agtttctttt tggttattcc ttgactgctt gattacagaa agaaatttat 2100
gaagctgtaa tggggatagt tatactgctt gttcttatga ttcatttctt ttgtgcagtt 2160
cttgggtgtag cttgccactt tcaccagcaa agttc 2195

```

<210> 2

<211> 61

<212> DNA

<213> Artificial sequence

<220>

<223> sense primer attB1

<400> 2

```

ggggacaagt ttgtacaaaa aagcaggett cacaatgta cgtcctgtag aaaccocaaac 60

```

c

61

072-GOS2-PROVL ST25.txt

<210> 3

<211> 52

<212> DNA

<213> Artificial sequence

<220>

<223> antisense primer attB2

<400> 3

ggggaccact ttgtacaaga aagctgggtt tggtgattca ttgtttgcct cc

52

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.